Polygenic Scores: are they a public health hazard?

Keith Baverstock
Department of Environmental and Biological Sciences
University of Eastern Finland
Kuopio Campus
Kuopio, Finland

Abstract

I argue here that polygenic scores are a public health hazard because the underlying methodology, genome wide association, from which they are derived, incorrectly assumes that the information encoded in the genomic DNA sequence is causal in terms of the cellular phenotype. This is not so when the cell is viewed from the perspective of a) fundamental physics, b) the protein chemistry that characterises the cellular cytoplasm and c) the fundamental requirement for evolution to yield unlimited species diversity.

The discipline of biophysics has not provided biology with the deep underpinning in physics, as, for example, atomic theory has provided for chemistry. In 1935 the physicist Max Delbrück defined genetics as "*.... a far-reaching, logically closed, strict science. It is quantitative without making use of a physical measurement system.*"[1] (Timoféeff-Ressovsky, Zimmer et al. 1935). Delbrück acknowledges that chemistry was transformed from a purely descriptive science by using a measurement system to establish a quantitative balance between reactants and products and, therefore, the conservation of mass. In this way causality in chemical reactions was established and mediating mechanisms subsequently discovered. Delbrück rejected the idea that biology could be similarly underpinned because chemistry depended on the given of a stable unchanging atom, whereas, living beings, the natural unit for quantitative analysis in genetics, are naturally changing and, therefore, genetics is independent of a measurement system (Timoféeff-Ressovsky, Zimmer et al. 1935) Genetics relates DNA sequences in the genotype, supposedly causally, to cellular phenotypic properties. It can say nothing about how cells work, or what is the physical nature of the crucial cellular phenotype. It is blind to the processes that would convert gene sequence to phenotypic properties and causality is, thus, only an article of faith, rather than something self-evident, as in a chemical reaction.

Can physics help us understand better how the cell works? Based on what a cell self-evidently is, a foundation in terms of basic physics, namely thermodynamics and complex dissipative system dynamics, has been proposed (Baverstock 2000, Annila and Baverstock 2014, Baverstock 2016). This, the so-called independent attractor (IA) model of the cell, predicts that there will not be a causal relationship between genotype and phenotype. This is because the output from the cell, in terms of its phenotype, is primarily epigenetic[2], rather than genetic. Furthermore, as plausible as things look, close examination of the essential informational link between genes and phenotypic properties, Crick's sequence hypothesis, is flawed (Baverstock 2019). Crick assumed that peptides, by *folding themselves* to the native protein structure, effectively translated sequence information from the gene into structural information

[1] "*Bekanntlich ist die Genetic eine weitgebend in sich logisch geschlossene, strenge Wissenschaft. Sie ist quantitaiv, ohne vom physicalischen Maßsystem Gebrauch zu machen.*"

[2] The term is used in its generic sense of "over and above genetics" and not in the context of chromatin/DNA marking (see below).

in the active proteins that inform the phenotype (Crick 1958). Later he appears to back away from this position, pointing to the hypothesis being based on expediency rather than any deep theoretical principle (Crick 1970). In fact, the rate at which peptides fold to proteins is so slow that very little native state protein forms in the cytoplasm, which is mainly populated by folding and unfolding peptides and mis-folded proteins (Baverstock 2019). That there is no contiguous flow of information from the DNA sequence to the phenotype is, thus, a second reason why the DNA sequence is not causal. Furthermore, the all-pervasive idea that the DNA sequence alone can determine cellular phenotype, as in the genetic regulatory network (GRN) model, is deeply flawed. The cell is not a Turing machine: reference by the formal syntactic system to a second source of information is required for completeness (in Gödel's sense) (Baverstock 2011). Additionally, the British geneticist and developmental biologist, Conrad Waddington, noted in 1968 that evolution could not have yielded unlimited diversity without a mutual interaction between environment and phenotype (Waddington 2008).

According to Paneth and Vermund, human molecular genetics has not yet contributed to measurable public health advances, despite unprecedented investments in resources since 1988. The National Human Genomics Research Institute at the National Institutes of Health has received at total of US$10 billion for the sequencing of the human genome (3 billion) and research to understand the role of the genome in human health (7 billion). This is, of course, a fraction of the total research budget disbursed worldwide over the past 30 years on molecular genetics and genomics. Furthermore, since the late 1950s eight Nobel Prizes have been awarded for research in molecular genetics compared to four for research yielding public health benefit. Nobel wished to reward work that benefited human kind as well as it being of high scientific merit (Paneth and Vermund 2018).

Due to the reduction in costs of genome sequencing, the technique of genome wide association (GWA) has been deployed over the past decade in a plethora of studies on patients with common diseases, such as depression, cancer, coronary artery disease, type II diabetes etc., as well as studies on complex traits, such as schizophrenia, human height, educational attainment and IQ. The results follow a remarkably similar pattern. For studies of tens, to tens of thousands, of patients, several to many

abnormal loci are detected with genome wide p-values < $5x10^{-8}$, each with very small effect and accounting for only a small fraction of the inheritable variance or risk. An iconic case in point is the 2014 study of Schizophrenia with some 37,000 subjects and 113,000 controls, which found 128 abnormalities, single nucleotide polymorphisms (SNPs), at 108 loci and accounted for 3.4%, or less, of the estimated heritable risk. (Schizophrenia Working Group of the Psychiatric Genomics 2014). Clearly, such results have no public health or clinical utility.

As stated above, there are three clear reasons why such GWA studies should be dismissed. In 1958 the American geneticist, David Nanney, drew a distinction between two modes of action in the cell – the template mechanism involving the decoding of the DNA to proteins, from which the phenotype was derived and *"... auxiliary mechanisms with different principles of operation ..... involved in determining which specificities are to be expressed in any particular cell.*" genetic and epigenetic systems as being on an equal footing, perhaps even suggesting that the epigenetic system might be dominant. He stresses the difficulty in assigning cellular features to one or the other system (Nanney 1958). The possibility, therefore, exists that the cell's output is, in fact, epigenetically and not genetically, regulated. If so, the gene products and not the genes, would be causally influencing the phenotype (Baverstock and Rönkkö 2008). This distinction between gene and gene product is important, because post-translational processes, especially peptide to protein folding, can modify the gene products independently of the originating gene's sequence.

The results from four radiobiological experiments published prior to 1993 (Luning, Frolen et al. 1976, Pampfer and Streffer 1988, Pampfer and Streffer 1989, Kadhim, Macdonald et al. 1992) defy explanation in terms of inherited phenotypic changes being encoded in DNA sequences, i.e., Nanney's 'genetic system': thus, they must be interpreted as consequences of epigenetic regulation of the cell (Schofield and Kondratowicz 2018). This can be explicitly understood in terms of the last of the above four experiments. The non-clonal chromosomal aberrations observed in this experiment cannot be the direct result of the action of radiation on the DNA, but rather must be the consequence of the response of the cell to the stress caused by the radiation (Baverstock 2000). The various responses of cells, programmed and un-programmed, to 'shocks', was the subject of Barbara McClintock's Nobel lecture

(McClintock 1984). The other three experiments can be similarly rationalised in terms of an epigenetic response to stress.

The phenomenon causing the non-clonal chromosomal aberrations was termed chromosomal, or genomic, instability (Kadhim, Macdonald et al. 1992), but this is a misnomer, as is clear from the experiment by Luning et al (Luning, Frolen et al. 1976). Surviving male mice from litters exhibiting increased incidences of intra-uterine death (IUD), fathered by a mouse with alpha-particle irradiated germ cells, without further irradiation, produced offspring with an increased yield of IUD, a dominant lethal mutation, when mated with unirradiated females. IUD should be lethal in early life, yet the male mice from a litter exhibiting increased IUD survived to adulthood to pass on a dominant lethal mutation. In this case irradiation of the father had modified the phenotype of his offspring. Therefore, what is termed genomic instability is, in fact, phenotypic instability according to the above evidence. An intervention by Delbrück at a genetic conference in Paris in 1949 shows that phenotypic transitions that do not involve modification of the DNA have long been a recognised phenomenon. In discussion following a paper presented by Sonneborn, Delbrück says: "*many systems in flux equilibrium are capable of several different equilibria under identical conditions. They can pass from one state of equilibrium to another under the influence of transient perturbations.*" (Delbruck 1949). Today, 'flux equilibrium' would be termed an 'attractor state', as proposed for the cellular phenotype in the IA model (Baverstock 2000, Baverstock and Rönkkö 2008).

Phenotypic instability can be understood in the context of the quasi-stable attractor state of a complex dissipative system (CDS), such as the cell is. In this respect the IA model underpins phenotypic instability based on attractor states of CDSs that represent the cellular phenotype. CDSs as complex as is the cell, harbour numerous potential attractor states (Baverstock 2011) and a stress induced loss of one attractor state implies, in most cases, the adoption of a variant attractor state and, therefore, a variant cellular phenotype. Phenotypic instability is, therefore, imposed upon the cell by the physical nature of the cellular phenotype.

That the phenotypic realisation of a genome is not fixed as a one-to-one relationship is clear from the fact that all the cells in an organism, regardless of their phenotypes,

have the same genomic sequence and that phenotypes as radically different as a caterpillar's is from that of its butterfly, are derived from a single DNA sequence. This is essentially what Nanney saw in 1958 when he pointed to the two complimentary systems in the cell (Nanney 1958). Today, that epigenetics is an essential component of cellular function is not doubted, but it is now most commonly regarded as being controlled by factors such as chromatin structure, acetylation of chromatin and methylation of DNA, rather than it being an integral component in cellular function as proposed by Nanney (Nicoglou and Merlin 2017).

If it were the case that the output of the cell, in terms of its phenotype, is via the genetic, rather than the epigenetic system, that is, the cell is seen as a 'machine' rather than a 'CDS', and subject to Newtonian physics, then, the vital route by which the information encoded in the genomic DNA sequence is assumed to be transformed to structural information on proteins, would be blocked by the slowness of the peptide-to-protein folding process (Baverstock 2019). It is also the case that the phenomenon termed 'genomic instability' cannot be understood in terms of a 'machine' model of the cell (CDS physics is required), or of a Mendelian inheritance process, or, as yet, in terms of chromatin/DNA marking (Schofield and Kondratowicz 2018).

The GRN model of the cell (Babu, Luscombe et al. 2004), upon which the GWA approach is based, relies on hard-wiring of the phenotype to the genotype through binding sites on proteins and recognition sites on DNA, without any external referents: the cell is viewed as a Turing machine (Baverstock 2011). The mathematical biologist, Robert Rosen, argues that a purely syntactic system such as a Turing machine, is incomplete (in Gödel's terms) without a semantic partner (Rosen 2000). Conrad Waddington, argued in 1968, in his "Paradigm for an Evolutionary Process" that the unlimited diversity in terms of species produced by evolution is only possible if phenotype can influence environment and environment can influence phenotype. Waddington reviews the work of those who laid the mathematical foundations of evolutionary theory[3] and identifies two neglected problems in formulating a theory of evolution, namely adaptation and speciation. He argues that these are key to evolution generating unbounded diversity. Random mutation alone,

[3] Ronald A Fisher, J B S Haldane and William Bateson

as a source of new variation, will not account for adaptation and without diversity of the environment through interactions with phenotypes, speciation would be limited (Waddington 2008).

Fundamental physics, protein chemistry, computational science and fundamental evolutionary theory all argue for the invalidity of the GWA methodology.

Notwithstanding decades of genetic research, the evidence today propels us to the conclusion that whatever genomic abnormalities GWA studies are measuring, they are not the cause of the trait under investigation: the genomic DNA does not directly influence the phenotype: they are independent of one another. Nanney's 'genetic system' transcribes and translates the genomic DNA database to provide the cytoplasm with inactive gene products (Nijhout 1990) and the phenotype processes those products to yield itself and to regulate the cell (Baverstock and Rönkkö 2008). In the special case of a single gene sequence yielding a single peptide that is utilised to produce a single trait, there is an association between the sequence and the trait. This is mostly the case for rare diseases such as Ehlers-Danlos syndrome (Parapia and Jackson 2008) and for simple traits, such as flower colour (Hellens, Moreau et al. 2010). However, rare inherited monogenic traits do not necessarily behave in the deterministic way predicted by classical genetics. For example, Krabbe disease is a rare monogenic condition caused by a deficiency in the enzyme galactocerebrosidase. In a screening trial of 1.9 million infants conducted over 8 years in New York City, 348 infants were referred for further follow-up using molecular analysis, based on the low levels of the enzyme detected. Of these two were lost to follow-up and five were confirmed to exhibit infantile Krabbe disease. However, the authors note that: "*Except for variants that clearly abolish protein function, such as the 30-kb deletion and other truncation, frameshift, nonsense, and splicing mutations, it is difficult to infer genotype–phenotype correlations because most variants are novel and/or rare.*" (Orsini, Saavedra-Matiz et al. 2016). Thus, the associations that characterise rare diseases[4] suggest that only a minority of mutations impact on phenotype: the majority are without effect.

[4] According to Genomics England some 7% of the population are affected by up to 8000 rare genetic conditions: https://www.genomicsengland.co.uk/understanding-genomics/rare-disease-genomics/

Common diseases and complex traits, for example, human height, where the trait is continuously distributed, present a different problem. As early as 1918 Fisher proposed that a continuously distributed trait would result from the combined effect of several genes, i.e., it would be polygenic. Fisher was a strong advocate of Mendel's laws of inheritance, *albeit* that he recognised that the result of Mendel's experiments had been fudged and that the probability of the results deviating as little as they did from the mean, was 1/4000 (Elston 2018). Mendel had emphasised the particulate nature of the units of inheritance as a means of discriminating between segregation and blending in the inheritance process. Polygenicity was an obvious solution to explain continuously distributed traits. However, as pointed out by the American geneticist, Richard Lewontin, experimental geneticists largely ignored complex traits, concentrating on the more easily measurable, but uninteresting, simple or monogenic traits (Lewontin 1974). This bias, it can be argued, has created an illusion, namely that genetics is a successful branch of biology. Of course, given the long-established phenomena of epistasis and the more recently uncovered phenomenon of moonlighting proteins(Gancedo and Flores 2008), difficulties might have been anticipated in relating, unequivocally, genotypes with several genes acting together to yield common or complex traits, to the phenotype. Also, a warning about the impact of alternative splicing of mRNA, on the prediction of phenotypes, was issued in 2000: the *DSCAM* gene found in *Drosophila* is theoretically capable of producing ~38,000 different proteins (Black 2000).

While traditional linkage studies were successful in identifying the single high-risk genes responsible for rare disease traits, where several genes, each with a low risk, were assumed to be involved in common disease traits, it was recognised that a different approach was needed (Tabor, Risch et al. 2002). Based on hypotheses concerning the biological plausibility of the involvement of specific genes in a trait, the 'candidate gene' approach was developed, in which the association of variants in those genes, with the trait, was determined statistically. These were termed "genetic association studies". However, it became clear that few if any of these candidate gene hypotheses were correct. For example, in 2012 it was clear that most reported candidate genes for general intelligence were false positives (Chabris, Hebert et al.

2012). More recently, a study with substantial statistical power, finds no support for 18 of the most prominently reported candidate genes, (including those that act through interaction with the environment) for depression (Border, Johnson et al. 2019). Genes are hypothesised to be candidate genes because they are plausibly associated with a trait that is more frequently seen in specific families. Linkage studies across the family members can identify regions of the genome where specific alleles are shared (Kwon and Goate 2000). Association of variants of these genes can then be sought in populations exhibiting the trait. That the candidate gene approach, based, as it is, on classical genetics, has failed, should have sounded a strong warning that perceptions of the role of genes in biology might be flawed. Instead of heeding this warning molecular genetic research focussed on the hypothesis-free approach of GWA.

As the technology to sequence the genome became cheaper, the GWA approach boomed with little *a priori* thought about the biological relevance of what was being measured in terms of SNPs. The approach was the reverse of 'hypothesis-driven': rather 'biological insights' were claimed from the knowledge about where in the genome variations were found (Schizophrenia Working Group of the Psychiatric Genomics 2014). For example, "*Associations were enriched among genes expressed in brain providing biological plausibility for the findings*". This hypothesis-after-the-results-are-known is called HARKing (Kerr 1998). The danger is that HARKing can be, and is, used to build 'castles in the air' by looking for common variant regions in different traits, related or not.

As already noted, across the domain of common disease and complex traits, GWA studies have produced a remarkably common picture, of several genes, each with very small effect, contributing to a small fraction of the assumed heritable variance. Attempts to increase the fraction of variance accounted for by expanding the number of study participants fails to focus attention on one, or a few, of those SNPs, but rather increases the number of SNPs identified for a marginal gain in the fraction of the variance accounted for. Therefore, the much sought after clinical utility of GWA studies *per se*, has proved elusive and the construction of so-called polygenic scores (PGSs) has been proposed as a means to sum-up the miniscule contribution of each SNP (Harlaar, Butcher et al. 2005, Plomin 2018). But according to the above considerations, GWA studies cannot produce meaningful data to be summed-up.

Barton et. al., in reviewing two GWA studies investigating the genetic basis for variation in human height, issue a strong caution. Variation in height could be environmentally controlled (by diet) or genetically controlled (inferred from twin/family studies) and subject to selection. The two studies independently confirm that selection plays no role in the gradient of human height observed north to south across Europe and that there is no genetic basis for such a variation. What is observed in studies claiming a genetic basis is due to the cumulative effects of biases in the population data base (GIANT) (Barton, Hermisson et al. 2019). However, Barton et al still maintain: "*that genetics plays a major role in height differences between individuals is not in doubt*". This conclusion is presumably based on twin/family studies, which only say that identical twin pairs are more likely to be of similar height than same sex fraternal twin pairs, thus height is an inherited trait and ergo, it is genetic. The psychologist Jay Joseph maintains that such studies are fatally confounded by environmental factors in any case (Joseph 2015). The idea that inheritance may not be mediated directly by genes lays outside of the closed methodology of genetics but not outside the domain of fundamental physics.

That GWA studies are detecting SNPs is not doubted: the question is, what is it about the participating patient/subject they are characterising? Kerminen et al[5] show that even with a genetically comparatively homogeneous population, such as that of Finland, PGSs calculated for common diseases (e.g., coronary artery disease) reveal the geographical stratification of the study population, rather than the trait. Barton et al warn that the papers they reviewed on the distribution of European heights "*demonstrate the potential for population structure to create spurious results, especially when using methods that rely on large numbers of small effects, such as polygenic scores. Caution is clearly needed when interpreting and using the results of such studies.*" (Barton, Hermisson et al. 2019).

In a comprehensive assessment of the role of GWA studies and PGSs in studies of cognitive ability and educational attainment, Richardson and Jones note that the GWA technique is prone to give rise to spurious correlations and to be primarily a measure of the underlying genetic population structure, in this case social class, rather than the

[5] https://www.biorxiv.org/content/10.1101/485441v1

trait. Correcting for this using socio-economic status is wholly inadequate (Richardson 2017, Richardson and Jones 2019). Although this analysis is directed at two specific traits, the conclusions are generally applicable. It is, therefore, clear from an empirical perspective that GWA studies are not measuring trait related abnormalities and from the arguments above the reason is that there is no causal basis for the influence of DNA sequence on phenotype: PGSs derived from patients and subjects have no predictive or explanatory utility in medicine.

There have been sceptics of the human genome sequencing enterprise from the outset. For example, in a scathing article in the New York Review of Books[6] in 1992, "The Dream of the Human Genome", Lewontin describes the molecular biological revolution of the mid-1900s, which kicked-off the sequencing enterprise, as having achieved "*a state of unchallenged orthodoxy*"; in 1994 in an essay published in the Guardian Newspaper,[7] I wrote: "*I fear though we can expect to learn from this great labour* [the sequencing of the human genome] *as much about how life 'works', or in the case of disease, does not work, as we can learn about how a telephone exchange works, from a telephone directory*", and in 2001, the day after the nearly completed human genome sequence was celebrated at the White House, Stephen Jay Gould writes in the New York Times[8], under the title: "Humbled by the Genome's Mystery*" "Human complexity cannot be generated by 30,000 genes* [it is now more like 20,000] *under the old view of life embodied in what geneticists literally called (admittedly with a sense of whimsy) their 'central dogma'*". On the other hand, in 1999, Francis Collins, leader of the Human Genome Project from 1993, and currently Director of the NIH, wrote that sequencing the human genome would lead: "*to previously unimaginable insights, and from there to the common good* [including] *a new understanding of genetic contributions to human disease and the development of rational strategies for minimizing or preventing disease phenotypes altogether*." (Collins 1999).

The reasons that Collins' optimism was misplaced is that human health, wellbeing and achievement are not based on an algorithm premised by a DNA base sequence,

[6] https://www.nybooks.com/articles/1992/05/28/the-dream-of-the-human-genome/
[7] https://www.newspapers.com/newspage/260739393/
[8] https://www.nytimes.com/2001/02/19/opinion/humbled-by-the-genome-s-mysteries.html

however sophisticatedly manipulated. Viewed from outside the logically closed methodology of genetics, it is clearly not the case that DNA "*tells us who we are*" as psychologist Robert Plomin claims in his recent book, "Blueprint" (Plomin 2018). Plomin claims that PGSs are a "*fortune-telling*" device, and indeed they are, in the same sense as are astrology and tea leaves. PGSs are the ultimate 'snake oil' based on measuring noise[9]: some of the most expensive snake oil in human history.

It is concerning that research on the development and application of PGSs is supported by the UK Medical Research Council[10], my former employer. It is totally unthinkable that the MRC would endorse a role for astrology in medical diagnosis and treatment, or in the development of social policy. Yet PGSs, taken at Plomin's face value, might find application in life changing situations. For example, in the education policy context, Plomin's specific interest: should a child's PGS, measured at birth for educational attainment, determine the extent of investment the state is prepared to make in his or her education? For several decades in the UK, entrance to grammar schools has been and is, based on intelligence type tests at age 11 years. Prior to 1965, with the introduction of comprehensive schools, the 80% who failed the tests were essentially abandoned by the education system: very few made it to university[11]. Further, Plomin maintains that "*predictions from polygenic scores are causal*" and that "*they will contribute to the demise of diagnosis*" (Plomin 2018). These remarks are made in the context of psychological/behavioural traits: however, PGSs are envisaged for application to common diseases and although some are cautious (Lewis and Vassos 2017), others have claimed success (Khera, Chaffin et al. 2018, Dichgans, Pulit et al. 2019). The results of Khera et. al. are, however, disputed (Curtis 2019). The dangers to individual patients posed by a false confidence in a low PGS, leading to serious disease being undiagnosed, or misplaced confidence in a high PGS, leading to, for example, biopsies, or even operations, for breast and prostate cancers, are all too real. From the perspective of public health, a continuing research focus on GWA studies as a major player in healthcare and the development of social policy, is highly counter-productive. The lost opportunity cost is unacceptable, given

[9] https://www.theatlantic.com/science/archive/2019/05/waste-1000-studies/589684/?utm_source=pocket-newtab

[10] https://kclpure.kcl.ac.uk/portal/en/organisations/mrc-centre-for-social-genetic--developmental-psychiatry(266adf4b-38eb-43f9-b777-58c20017d9c4).html

[11] https://en.wikipedia.org/wiki/Secondary_modern_school

the resources that are being devoted to research based on a fundamentally flawed methodology.

The answer to the question in the title is, therefore, ‘yes’. PGSs, the ultimate product of what is widely known as the ‘genomics revolution’, are a dangerous delusion. There is, therefore, an urgent need for a debate on how to better understand the causes of common diseases and complex traits: I would suggest that the role of phenotypic (formerly known as genotypic) instability would be a good topic upon which to refocus resources, but clearly the impact of the environment on health must also be a priority. But perhaps the most important lesson we can take from the last 60 years of genetic research is a point made by Robert Rosen in his 1991 book, “Life Itself”: “*Pleasing as you may think the view from the top of the skyscraper is, it is wise to visit the basement occasionally to check the foundations*.” (Rosen 1991)

Acknowledgements.

Comments by Hans Edelmann, Mike Jones, Jay Joseph, Steve Pittelli and Ken Richardson are gratefully acknowledged, as is the enlightenment derived from the discussions with the Critical Genetics Forum. Valuable suggestions by a reviewer are also gratefully acknowledged.

Conflicts of interest: there are no conflicts of interest.

References

Annila, A. and K. Baverstock (2014). "Genes without prominence: a reappraisal of the foundations of biology." J R Soc Interface **11**(94): 20131017.

Babu, M. M., N. M. Luscombe, L. Aravind, M. Gerstein and S. A. Teichmann (2004). "Structure and evolution of transcriptional regulatory networks." Current Opinion in Structural Biology **14**(3): 283-291.

Barton, N., J. Hermisson and M. Nordborg (2019). "Why structure matters." Elife **8**.

Baverstock, K. (2000). "Radiation-induced genomic instability: a paradigm-breaking phenomenon and its relevance to environmentally induced cancer." Mutation Research **454**(1-2): 89-109.

Baverstock, K. (2011). "A comparison of two cell regulatory models entailing high dimensional attractors representing phenotype." Prog Biophys Mol Biol **106**(2): 443-449.

Baverstock, K. (2016). "Genes without prominence." Inference: International Review of Science **2(2)**.

Baverstock, K. (2019). "Crick's sequence hypothesis: a review." Communicative and Integrative Biology **12**(1): 59-64.

Baverstock, K. and M. Rönkkö (2008). "Epigenetic regulation of the mammalian cell." PloS One **3**(6): e2290.

Black, D. L. (2000). "Protein diversity from alternative splicing: a challenge for bioinformatics and post-genome biology." Cell **103**(3): 367-370.

Border, R., E. C. Johnson, L. M. Evans, A. Smolen, N. Berley, P. F. Sullivan and M. C. Keller (2019). "No Support for Historical Candidate Gene or Candidate Gene-by-Interaction Hypotheses for Major Depression Across Multiple Large Samples." Am J Psychiatry: appiajp201818070881.

Chabris, C. F., B. M. Hebert, D. J. Benjamin, J. Beauchamp, D. Cesarini, M. van der Loos, M. Johannesson, P. K. Magnusson, P. Lichtenstein, C. S. Atwood, J. Freese, T. S. Hauser, R. M. Hauser, N. Christakis and D. Laibson (2012). "Most reported genetic associations with general intelligence are probably false positives." Psychol Sci **23**(11): 1314-1323.

Collins, F. S. (1999). "Shattuck lecture--medical and societal consequences of the Human Genome Project." N Engl J Med **341**(1): 28-37.

Crick, F. (1958). "On protein synthesis." Symp Soc Exp Biol **12**: 138-163.

Crick, F. (1970). "Central dogma of molecular biology." Nature **227**(5258): 561-563.

Curtis, D. (2019). "Clinical relevance of genome-wide polygenic score may be less than claimed." Ann Hum Genet **83**(4): 274-277.

Delbruck, M. (1949). "translation of discussion following a paper by Sonneborn, T. M. and Beale, G.H." Unites Biologiques Douees de Continute Genetique: 33 - 35.

Dichgans, M., S. L. Pulit and J. Rosand (2019). "Stroke genetics: discovery, biology, and clinical applications." Lancet Neurol **18**(6): 597-599.

Elston, R. C. (2018). "Fisher's influence on me." Genet Epidemiol **42**(8): 849-853.

Gancedo, C. and C. L. Flores (2008). "Moonlighting proteins in yeasts." Microbiol Mol Biol Rev **72**(1): 197-210, .

Harlaar, N., L. M. Butcher, E. Meaburn, P. Sham, I. W. Craig and R. Plomin (2005). "A behavioural genomic analysis of DNA markers associated with general cognitive ability in 7-year-olds." J Child Psychol Psychiatry **46**(10): 1097-1107.

Hellens, R. P., C. Moreau, K. Lin-Wang, K. E. Schwinn, S. J. Thomson, M. W. Fiers, T. J. Frew, S. R. Murray, J. M. Hofer, J. M. Jacobs, K. M. Davies, A. C. Allan, A.

Bendahmane, C. J. Coyne, G. M. Timmerman-Vaughan and T. H. Ellis (2010). "Identification of Mendel's white flower character." PLoS One **5**(10): e13230.
Joseph, J. (2015). The trouble with twin studies : a reassessment of twin research in the social and behavioral sciences. New York, NY, Routledge, Taylor & Francis Group.
Kadhim, M. A., D. A. Macdonald, D. T. Goodhead, S. A. Lorimore, S. J. Marsden and E. G. Wright (1992). "Transmission of chromosomal instability after plutonium alpha-particle irradiation." Nature **355**(6362): 738-740.
Kerr, N. L. (1998). "HARKing: hypothesizing after the results are known." Pers Soc Psychol Rev **2**(3): 196-217.
Khera, A. V., M. Chaffin, K. G. Aragam, M. E. Haas, C. Roselli, S. H. Choi, P. Natarajan, E. S. Lander, S. A. Lubitz, P. T. Ellinor and S. Kathiresan (2018). "Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations." Nat Genet **50**(9): 1219-1224.
Kwon, J. M. and A. M. Goate (2000). "The candidate gene approach." Alcohol Res Health **24**(3): 164-168.
Lewis, C. M. and E. Vassos (2017). "Prospects for using risk scores in polygenic medicine." Genome Med **9**(1): 96.
Lewontin, R. C. (1974). The genetic basis of evolutionary change. New York, Columbia University Press.
Luning, K. G., H. Frolen and A. Nilsson (1976). "Genetic effects of 239Pu salt injections in male mice." Mutation Research **34**(3): 539-542.
McClintock, B. (1984). "The significance of responses of the genome to challenge." Science **226**(4676): 792-801.
Nanney, D. L. (1958). "Epigenetic Control Systems." Proc Natl Acad Sci U S A **44**(7): 712-717.
Nicoglou, A. and F. Merlin (2017). "Epigenetics: A way to bridge the gap between biological fields." Stud Hist Philos Biol Biomed Sci **66**: 73-82.
Nijhout, H. F. (1990). "Metaphors and the role of genes in development." Bioessays **12**(9): 441-446.
Orsini, J. J., C. A. Saavedra-Matiz, M. H. Gelb and M. Caggana (2016). "Newborn screening for Krabbe's disease." J Neurosci Res **94**(11): 1063-1075.
Pampfer, S. and C. Streffer (1988). "Prenatal death and malformations after irradiation of mouse zygotes with neutrons or X-rays." Teratology **37**(6): 599-607.
Pampfer, S. and C. Streffer (1989). "Increased chromosome aberration levels in cells from mouse fetuses after zygote X-irradiation." Int J Radiat Biol **55**(1): 85-92.
Paneth, N. and S. H. Vermund (2018). "Human Molecular Genetics Has Not Yet Contributed to Measurable Public Health Advances." Perspect Biol Med **61**(4): 537-549.
Parapia, L. A. and C. Jackson (2008). "Ehlers-Danlos syndrome--a historical review." Br J Haematol **141**(1): 32-35.
Plomin, R. (2018). Blueprint : how DNA makes us who we are. Cambridge, MA, The MIT Press.
Richardson, K. (2017). "GWAS and cognitive abilities: Why correlations are inevitable and meaningless: GWA studies to identify genetic factors for educational achievements largely ignore underlying social structures and dynamics." EMBO Rep **18**(8): 1279-1283.
Richardson, K. and C. Jones (2019). "Why genome-wide associations with cognative ability measures are probably spurious." New Ideas in Psychology **55**: 35-41.

Rosen, R. (1991). Life Itself: a Comprehensive Inquiry into the Nature, Origin and Fabrication of Life. New York, Columbia University Press.
Rosen, R. (2000). Essays on life itself. New York, Columbia University Press.
Schizophrenia Working Group of the Psychiatric Genomics, C. (2014). "Biological insights from 108 schizophrenia-associated genetic loci." Nature **511**(7510): 421-427.
Schofield, P. N. and M. Kondratowicz (2018). "Evolving paradigms for the biological response to low dose ionizing radiation; the role of epigenetics." Int J Radiat Biol **94**(8): 769-781.
Tabor, H. K., N. J. Risch and R. M. Myers (2002). "Candidate-gene approaches for studying complex genetic traits: practical considerations." Nat Rev Genet **3**(5): 391-397.
Timoféeff-Ressovsky, N. W., K. G. Zimmer and M. Delbrück (1935). "Über die Natur der Genmutation und der Genstruktur." Nachrichten der Biologischen Gesellschaft für Wissenschaft, Göttingen **1**: 189 - 241.
Waddington, C. H. (2008). "Paradigm for an Evolutionary Process." Biological Theory **3**(3): 258-266.